

Molecular Phylogenetics of *Candida albicans*^{▽†}

Frank C. Odds,^{1*} Marie-Elisabeth Bougnoux,^{2,3} Duncan J. Shaw,¹ Judith M. Bain,¹
Amanda D. Davidson,¹ Dorothee Diogo,² Mette D. Jacobsen,¹ Maud Lecomte,²
Shu-Ying Li,⁴ Arianna Tavanti,⁵ Martin C. J. Maiden,⁶
Neil A. R. Gow,¹ and Christophe d'Enfert²

Aberdeen Fungal Group, School of Medical Sciences, Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, United Kingdom¹; Unité Biologie et Pathogénicité Fongiques, INRA USC 2019, Institut Pasteur, Paris, France²; Laboratoire de Parasitologie-Mycologie, Service de Microbiologie, Hôpital Necker-Enfants Malades, Université Paris 5-René Descartes, Faculté de Médecine, Paris, France³; Laboratory for Mycology, Division of Research and Diagnostics, Center for Disease Control, Taipei, Taiwan⁴; Dipartimento di Biologia, Via San Zeno 35-39, Università di Pisa, 56127 Pisa, Italy⁵; and The Peter Medawar Building for Pathogen Research and Department of Zoology, University of Oxford, OX1 3SY Oxford, United Kingdom⁶

Received 6 February 2007/Accepted 23 March 2007

We analyzed data on multilocus sequence typing (MLST), ABC typing, mating type-like locus (*MAT*) status, and antifungal susceptibility for a panel of 1,391 *Candida albicans* isolates. Almost all (96.7%) of the isolates could be assigned by MLST to one of 17 clades. eBURST analysis revealed 53 clonal clusters. Diploid sequence type 69 was the most common MLST strain type and the founder of the largest clonal cluster, and examples were found among isolates from all parts of the world. ABC types and geographical origins showed statistically significant variations among clades by univariate analysis of variance, but anatomical source and antifungal susceptibility data were not significantly associated. A separate analysis limited to European isolates, thereby minimizing geographical effects, showed significant differences in the proportions of isolates from blood, commensal carriage, and superficial infections among the five most populous clades. The proportion of isolates with low antifungal susceptibility was highest for *MAT* homozygous *a/a* types and then α/α types and was lowest for heterozygous *a/\alpha* types. The tree of clades defined by MLST was not congruent with trees generated from the individual gene fragments sequenced, implying a separate evolutionary history for each fragment. Analysis of nucleic acid variation among loci and within loci supported recombination. Computational haplotype analysis showed a high frequency of recombination events, suggesting that isolates had mixed evolutionary histories resembling those of a sexually reproducing species.

Candida albicans is the most frequently encountered among the *Candida* species associated with humans as human commensals and opportunistic pathogens (10, 46). Over many years, this pleomorphic yeast has attracted much research attention; it was the first human fungal pathogen for which the full genome sequence was determined (30). *C. albicans* is a diploid species without sexual morphology and with substantial heterozygosity (48). Its predominant mode of reproduction is clonal (1, 23, 37, 41, 52), but it can occasionally undergo recombination (1, 23, 65), mitotic crossing over (48), and ploidy changes based on chromosome loss and reformation (58, 72). These contribute to the genomic microvariation that has been documented for multiple *C. albicans* isolates from single patients (8, 37, 47, 50, 56).

C. albicans possesses homologs of the mating type genes in *Saccharomyces cerevisiae* (25) and can be induced to undergo a mating process in which nuclei from two diploid cells fuse to form a tetraploid (29, 36). However, the fused nuclei randomly lose chromosomes to return to the diploid state rather than undergoing a true process of meiosis (3, 4), probably because

C. albicans lacks some of the genes required for meiosis in other yeasts (70). The mating process occurs between diploid strains homozygous for the mating types *a/a* and α/α , which arise spontaneously in a minority of clinical isolates (35, 38, 47, 63). However, mating usually occurs only after strains of opposite mating types have been induced to undergo a transition from the commonly observed, spheroidal, white cell form to the elongated, papule-surfaced, opaque cell form. This form survives at room temperature rather than body temperature and appears to be primed to undergo a mating process (34, 44, 69). Although the opaque cell form of *C. albicans* is associated in the laboratory with temperatures lower than 37°C, coinjection of genetically marked white cells of opposite mating types intravenously into mice led to mating *in vivo* (26), which suggests that the white-opaque switch might also occur in injected tissues. Recent evidence suggests that pheromone signals from rarely occurring opaque cells may induce biofilm formation among white cells of opposite mating types, which then enhances the close contact needed to induce mating *in vivo* (16).

Haplotype analyses of *C. albicans* strains show that while the dominant mode of reproduction in the species is clonal, recombination events also occur (65). Moreover, the spontaneous occasional occurrence of homozygous *a/a* or α/α isolates in the course of longitudinal sequences of *a/\alpha* isolates from single individuals (47) strongly suggests that *C. albicans* does indeed undergo steps known to be associated with mating and white-

* Corresponding author. Mailing address: Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, United Kingdom. Phone and fax: 44 1224 555828. E-mail: f.odds@abdn.ac.uk.

† Supplemental material for this article may be found at <http://ec.asm.org/>.

[▽] Published ahead of print on 6 April 2007.

opaque switching in colonized and infected patients. Full and partial chromosomal loss and replacement are well demonstrated phenomena for *C. albicans* (12, 31, 58, 72), and microevolution (also sometimes called microevolution) has been demonstrated by various strain typing approaches (8, 37, 47, 56). We therefore previously suggested that the population of cells colonizing or infecting a particular site in a patient comprises a mixture of cells with nearly identical genomes (47).

Multilocus sequence typing (MLST) has gained widespread acceptance as a tool for exploring strain-level differences in microbial species; it has been used for typing of many pathogenic bacteria (40), and systems have now been devised for several pathogenic fungi, including *C. albicans* (6, 7, 9), *Candida glabrata* (18), *Candida krusei* (28), and *Candida tropicalis* (64). MLST combines a high discriminatory index with exceptional portability since MLST data can be accessed and updated on the Internet (linked via <http://www.mlst.net/> or <http://pubmlst.org/>). Sequence data from multiple loci can be used not only to analyze clonality and recombination for strain populations within a species but also to provide evidence for species-level differentiation (67). A population genetic analysis of *C. albicans* MLST data based on 416 isolates from separate sources (63) showed that the species could be divided into several clades and that clades differed in the proportions of isolates they included from different geographical and anatomical sources. Since our *C. albicans* MLST database has now grown to include well above 1,000 isolates, with a more equitable distribution of isolates from different geographical sources, we undertook a fresh analysis of the data to provide a more robust reference basis for definition of MLST clades, to characterize further the prevalence of genetic recombination events, and to explore further the mechanisms underlying microdiversity in *C. albicans*.

We show here that 97% of the isolates could be assigned to one of 17 clades; that the proportions of A, B, and C genotypes, defined by the presence or absence of an intron in the ribosomal DNA region, differed significantly among clades; that clades were enriched with isolates from particular geographical areas; that the five most populous clades differed significantly in their proportions of commensal and pathogenic isolates; and that phylogenies of the seven fragments sequenced were not congruent, implying independent evolution of the fragments. A computational haplotype analysis of the sequences indicated that recombination events have led to isolates with mixed evolutionary histories, resembling the picture seen with a sexually reproducing species.

MATERIALS AND METHODS

***C. albicans* isolates.** A panel of 1,391 *C. albicans* isolates was assembled from MLST results for seven gene loci, namely, *AAT1a*, *ADP1*, *ACC1*, *MP1b*, *SYA1*, *VPS13*, and *ZWF1b*, contributed to the central database (<http://test1.mlst.net/>). Results of MLST are retained in the database in the form of genotype numbers, which define unique sequences for pairs of alleles, and diploid strain types (DSTs), which define unique combinations of genotypes. In this study, each isolate came from a single patient, with the exception of 53 isolates. In these instances, different unique DSTs were encountered among separate samples from the same patient, and the data were included to ensure that all known DSTs were analyzed. The isolate panel, which expanded the previously published set of 416 single-source isolates, included sets from sources in Asia, Africa, and South America in addition to the many isolates received from European countries and North America (Table 1). Forty-four different countries from most parts of the globe were therefore represented in the panel (Table 1). England/Wales and

Scotland were analyzed as separate geographical sources because of the very large sizes of the data sets available from these countries. Full details of the isolates are given in Table S1 in the supplemental material.

MLST, ABC typing, and MAT status determination. The methodology for determining DSTs by MLST has been described previously (9, 11). For 1,338 isolates, the ABC type was also determined, based on the presence or absence of an intron in DNA sequences encoding rRNA (42, 43). For 1,294 isolates, the mating type, i.e., *a/α*, *a/a*, or *α/α*, was determined by PCR (66). Susceptibility testing, performed on 658 of the isolates, was done by EUCAST methodology (14, 15). Isolates were defined as having reduced susceptibility to an agent, on the basis of Clinical Laboratory Standards Institute breakpoints (45), when the MICs of fluconazole, itraconazole, voriconazole, and flucytosine were $>8 \mu\text{g/ml}$, $>0.13 \mu\text{g/ml}$, $>1 \mu\text{g/ml}$, and $>4 \mu\text{g/ml}$, respectively.

Analysis of MLST data. The MLST results for single nucleotide polymorphisms (SNPs) in the seven sequenced loci were concatenated into a single sequence and converted with proprietary software to represent each SNP as two bases, as previously described (63, 64). This procedure generated sequences that could be handled by MEGA software, which cannot analyze heterozygous code data for nucleotide *P* distances. For any pair of diploid isolates, the result at each SNP can be homozygous and identical between the isolates, heterozygous and identical, homozygous and different, or heterozygous in one isolate and homozygous in another. For example, the sequencing result (in IUPAC single-letter code) for a given SNP across a set of strains might appear as A, T, or W (A or T). Concatenated data for each SNP in the seven *C. albicans* genes were therefore rewritten twice for homozygous (A, C, G, or T) data or once each for the two component bases for heterozygous (K, M, R, S, W, or Y) data. This procedure was the functional equivalent of scoring a pair of results with a 1 for homozygous or heterozygous identical data, 0 for homozygous different data, and 0.5 when one polymorphic site had a heterozygous result and the other had a homozygous result and then creating a difference matrix. The concatenated sequences were used to generate a single-linkage dendrogram for the 1,391 isolates by the unweighted-pair group method using average linkages (UPGMA), determined by *P* distances, as implemented by MEGA3 software (32). Clades were numbered to be consistent with an earlier publication on population structures (63). Clonal clusters of isolates that differed in sequence at only one of the seven loci were determined by eBURST, version 3.0 (21, 62). Analysis of variance (ANOVA) was done with the general linear model module of SPSS, version 14.0.

The congruence of maximum parsimony trees generated for each of the seven gene fragments sequenced for MLST was determined by an approach previously described for bacterial MLST data (20, 24). An MLST data set was chosen for 33 isolates that represented the diversity of the whole panel of isolates. The isolates were examples of the putative founding DSTs for 30/33 eBURST clusters which contained three or more DSTs and for which a founding DST was designated by the software. For three further eBURST clusters with multiple candidates as founding DSTs, a DST was chosen at random. Congruence tests were done by the method of Holmes et al. (24), with the software package PAUP*. For each gene, the optimum maximum likelihood (ML) tree was found. The data for each gene were then fitted in turn to the ML trees for each of the other six genes, and an ML score was calculated. Finally, the data for each gene were fitted to 200 trees of random topology, and ML scores were calculated.

The seven genes used for MLST analysis each showed multiple SNPs. Fast-Phase analysis of the SNP data generated a list of all haplotypes within the 1,391-isolate sample. If an isolate was heterozygous for zero or one SNP in a particular gene, then the SNP haplotypes for that gene could be determined unambiguously. If there were two or more heterozygous SNPs, the software employed a probabilistic approach to determine the likeliest haplotypes. If there was no recombination between haplotypes, nor any recurrent or reverse mutation, then *n* SNPs were expected to result in *n* + 1 haplotypes.

Population genetic tests were carried out by using the Arlequin package, version 3.01 (19). Haplotypes were inferred from MLST genotype data by Fast-Phase (54). Perl scripts were used for data simulation and for detection of putative recombinant haplotypes (see Table S2 in the supplemental material).

RESULTS

MLST of 1,391 *C. albicans* isolates resulted in delineation of 1,005 separate DSTs (for details, see Table S1 in the supplemental material). DST 69 was the most commonly encountered strain type: 65 isolates were type 69, and examples of this type were obtained from geographical sources in all parts of the world. The following six other DSTs were represented by

TABLE 1. Summary of panel of 1,391 *C. albicans* isolates used in this study

Geographical origin	No. of isolates obtained from source ^a								Total no. of isolates
	Blood	Oral cavity ^b	Vagina or vulva	Feces ^c	Urine	Other site	Animals	Unknown	
Argentina	4	0	0	0	0	0	0	0	4
Australia	10	1	0	0	4	4	0	0	19
Belgium	0	41	5	29	0	17	11	0	103
Brazil	12	6	0	0	0	0	0	2	20
Canada	2	23	0	0	0	0	0	3	28
Chile	4	0	0	0	0	0	0	0	4
China	13	0	0	0	0	0	0	0	13
Colombia	7	2	1	0	5	6	0	0	21
Czech Republic	4	0	0	0	0	0	0	0	4
Ecuador	4	0	0	0	0	0	0	0	4
England	100	84	49	10	2	14	0	12	271
Fiji	0	2	0	0	1	2	0	0	5
Finland	4	0	0	0	0	0	0	0	4
France	42	10	3	8	46	6	5	3	123
Germany	0	12	0	0	0	1	0	0	13
Hong Kong	44	0	0	0	0	0	0	0	44
Hungary	4	0	0	0	0	0	0	0	4
Ireland	3	0	0	0	0	0	0	0	3
Israel	1	0	2	0	0	0	0	0	3
Italy	4	5	0	0	0	0	0	0	9
Japan	27	0	32	0	0	0	0	0	59
Korea	0	1	0	0	0	0	0	0	1
Malaysia	0	0	0	0	3	1	0	0	4
Mexico	13	0	0	0	0	0	0	0	13
Netherlands	5	1	0	0	0	0	0	0	6
New Zealand	0	2	0	1	1	3	0	0	7
Nigeria	0	9	0	0	0	0	0	0	9
Peru	1	0	0	0	0	1	0	0	2
Poland	4	0	0	0	0	0	0	0	4
Portugal	12	0	0	0	0	0	0	0	12
Russia	2	0	0	0	0	0	0	0	2
Rwanda	2	14	0	0	0	1	0	0	17
El Salvador	0	0	0	0	0	0	0	1	1
Saudia Arabia	0	0	0	0	0	4	0	0	4
Scotland	147	76	0	0	1	2	0	0	226
Slovakia	2	0	0	0	0	0	0	0	2
South Africa	34	0	1	0	0	0	0	1	36
Spain	4	0	4	0	0	6	0	0	14
Switzerland	2	4	0	0	0	0	0	0	6
Taiwan	2	20	0	10	8	5	0	0	45
Turkey	8	0	0	0	0	0	0	0	8
United States	36	17	24	2	5	5	3	81	173
Venezuela	12	0	0	0	0	0	0	0	12
Zaire	0	0	0	0	0	0	2	0	2
Unknown	0	2	9	0	0	0	0	16	27
Total	575	332	130	60	76	77	21	120	1,391

^a Information regarding the source was unavailable for 17 isolates.

^b Includes sputa, throat swabs, and other samples from the oropharyngeal region.

^c Includes anal and rectal swab samples.

10 or more isolates: DSTs 37 (10 isolates), 79 (11 isolates), 155 (13 isolates), 182 (13 isolates), 277 (10 isolates), and 538 (10 isolates). None of these DSTs showed geographical limitation, either. The results indicate a discriminatory power for MLST of 0.996, calculated according to Hunter's formula (27).

Designation of clades. eBURST analysis of the MLST data assigned the 1,391 *C. albicans* isolates to 53 clonal clusters of DSTs differing from each other by a single genotype and to 368 singleton DSTs. Details of the clonal clusters that included more than three DSTs are summarized in Table 2. Clusters 24 to 33 each comprised three DSTs, and clusters 34 to 53 each comprised two DSTs. The largest clonal cluster, cluster 1,

based on DST 69 as the putative founding type, is illustrated in Fig. 1. Its complex structure was typical for the 10 largest clonal clusters and suggests a considerable level of recombination events among the DSTs. A highly clonal population would be expected to generate noncomplex clusters consisting mainly of a single DST.

Figure 2 illustrates the UPGMA dendrogram based on MLST data for 1,391 isolates of *C. albicans*. A full-sized table showing the relationships between the dendrogram, DSTs, eBURST clonal clusters, and other properties is provided as Table S3 in the supplemental material. To delimit clusters of closely related strain types, a cutoff *P* distance of 0.04 was

TABLE 2. Summary of data for eBURST clonal clusters with four or more DSTs

Clonal cluster	No. of isolates	No. of DSTs	Predicted founding DST ^a
1	355	186	69
2	124	84	124
3	135	75	155
4	58	47	345
5	54	34	881
6	32	20	735
7	22	16	300
8	22	15	766
9	19	13	173
10	9	9	840
11	17	7	167
12	10	7	90
13	6	6	378
14	9	6	305
15	7	6	195
16	10	6	410
17	6	5	322
18	5	5	—
19	7	4	366
20	5	4	—
21	5	4	935
22	5	4	—
23	4	4	380

^a —, no DST could be determined as the putative founding type.

chosen because this cutoff separated clusters that contained known examples of isolates in clades II and SA, as previously determined by DNA fingerprinting with the moderately repetitive probe Ca3 (61). On the basis of this cutoff value, 17 clades were defined (Fig. 2). Isolates that did not cluster with others within the 0.04 cutoff and clusters containing fewer than 10 isolates were labeled singletons. A total of 44 isolates received the singleton designation, leaving 1,347 isolates (96.8% of the panel) assigned to clades. The main features of the clades are summarized in Table 3. Although in the dendrogram (Fig. 2) the separation of the small clades 7 and 14 appears to be a borderline case for inclusion of all isolates in a single clade, the data in Table 3 indicated substantial differences in both the geographical origins and the ABC type distributions between the two clades, so we applied the 0.04 cutoff strictly and retained the separation for clades 7 and 14.

Clade 1, the largest clade, accounted for one-third (33.6%) of the 1,391 isolates tested. It contained all of the isolates in eBURST clonal cluster 1, with DST 69 as the most common individual strain type within the clade. Clades 2 to 4 were the next largest clades, with 34.0% of the 1,391 isolates being split among them. The four largest clades, clades 1 to 4, each con-

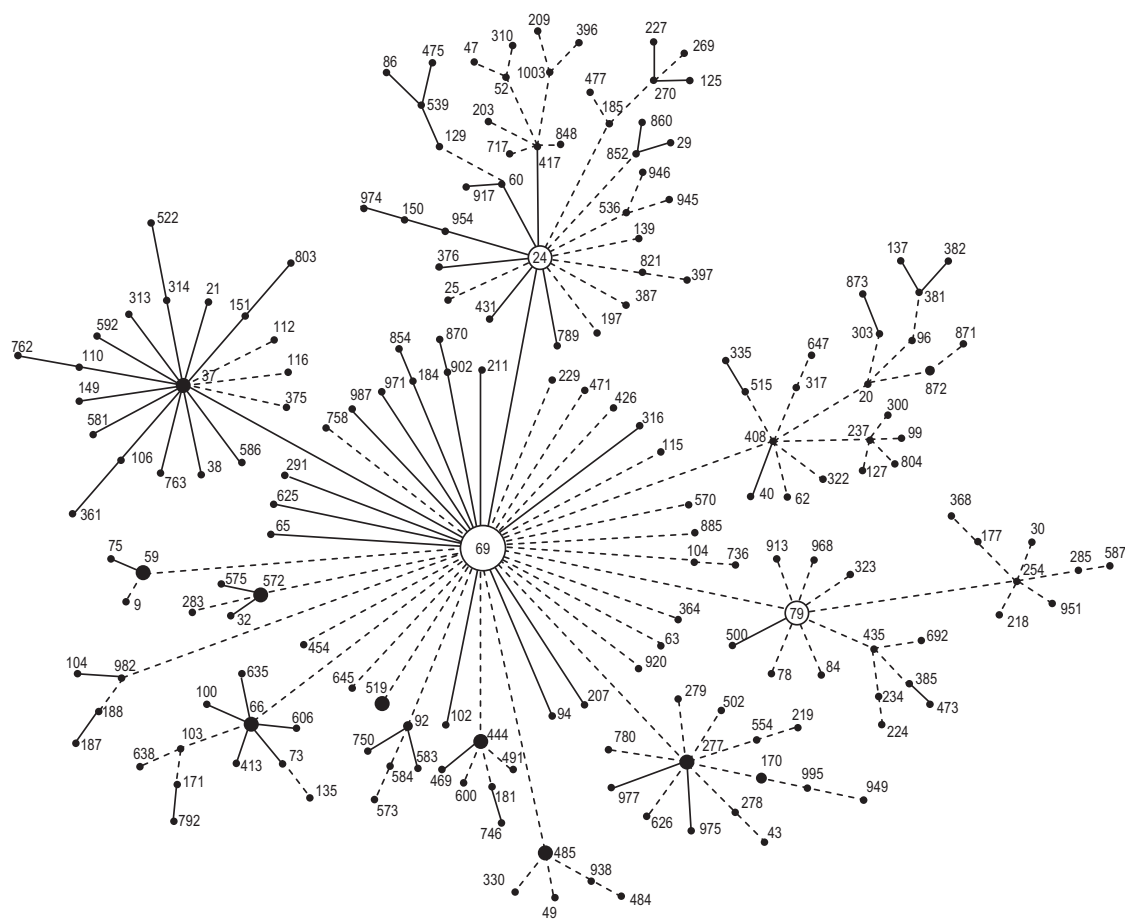


FIG. 1. eBURST clonal cluster 1. The illustration shows the figure generated by eBURST 3.0 software, in which DSTs are linked when they differ in the sequence of just one of the seven gene fragments used for MLST. The lengths of lines are not significant. The lines were redrawn to indicate differences between DSTs that result from a loss of heterozygosity from the putative parent to the putative derived isolate (dashed lines) or from mutation, gene exchange, or another mechanism (solid lines).

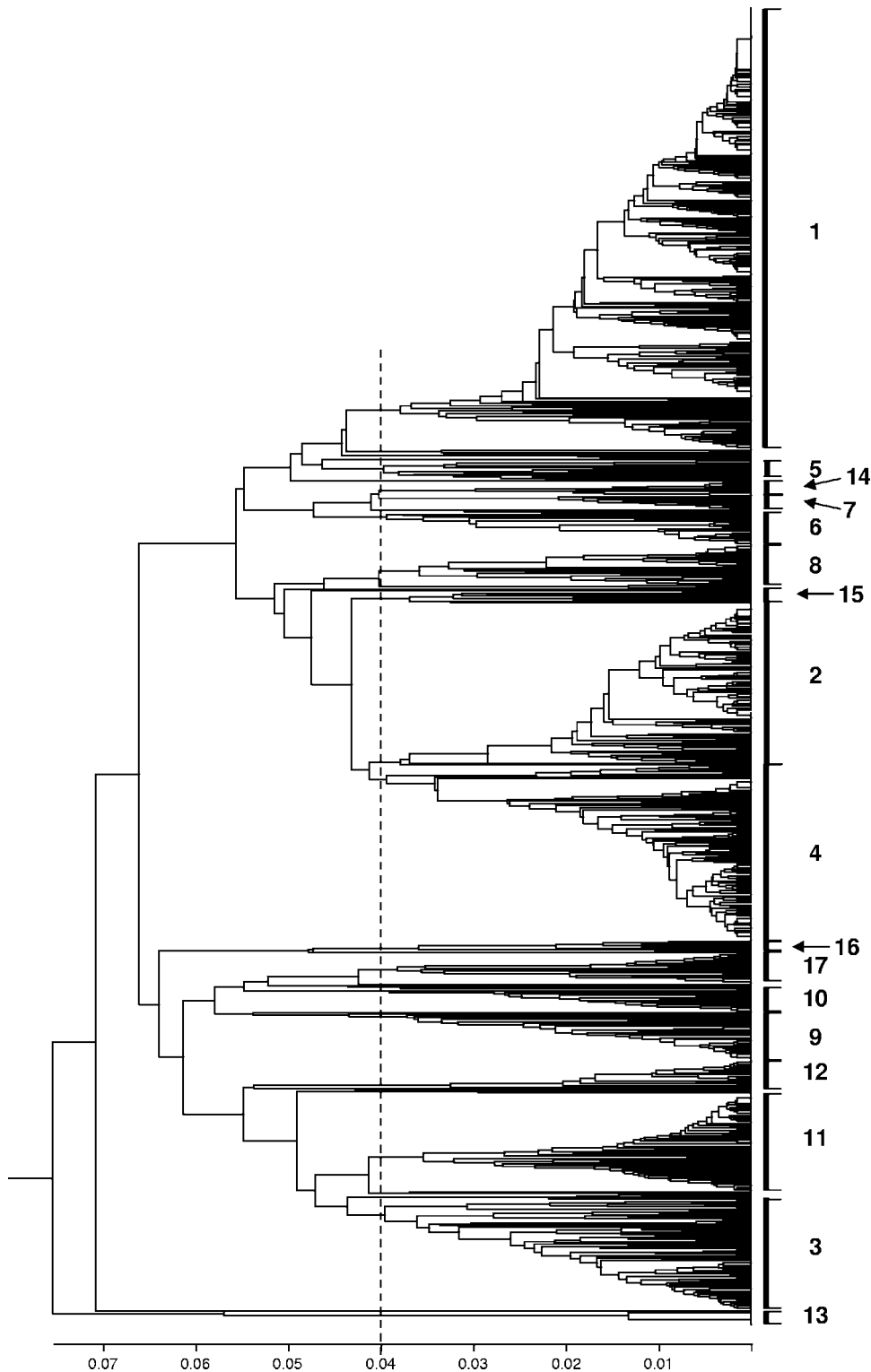


FIG. 2. UPGMA dendrogram showing P distances for 1,391 *C. albicans* isolates from single sources, or for multiples from single sources that represented unique DSTs, typed by MLST. The dashed line shows the cutoff at a P distance of 0.04 used to delineate clusters designated as clades. The number and vertical bar to the right of each clade indicate the clade designation and relative size, respectively.

tained five reference isolates previously typed by DNA fingerprinting (61) as belonging to clades I, II, III, and SA. Clades 1 to 4 thus show full correspondence with the previously delimited clades I, II, III, and SA, respectively. Among five reference

isolates assigned to clade E by DNA fingerprinting, two clustered with clade 4 isolates and three clustered with clade 11 isolates. Clade 11, previously represented by only nine isolates (63), emerged in the present study as the fifth most populous

TABLE 3. ABC types, *MTL* types, geographical origins, anatomical origins, dates of isolation, and antifungal susceptibilities of *C. albicans* isolates in 17 clades plus singleton isolates

Characteristic	n	No. (%) of isolates																		
		Clade 1	Clade 2	Clade 3	Clade 4	Clade 5	Clade 6	Clade 7	Clade 8	Clade 9	Clade 10	Clade 11	Clade 12	Clade 13	Clade 14	Clade 15	Clade 16	Clade 17	Singletons	
All isolates ABC type	1,391	467 (33.6)	171 (12.3)	117 (8.4)	185 (13.3)	20 (1.4)	36 (2.6)	14 (1.0)	45 (3.2)	49 (3.5)	24 (1.7)	104 (7.5)	29 (2.1)	14 (1.0)	16 (1.2)	13 (0.9)	11 (0.8)	32 (2.3)	44 (3.2)	
	857	418 (48.8)	149 (17.4)	21 (2.5)	47 (5.5)	3 (0.4)	12 (1.4)	14 (1.6)	22 (2.6)	36 (4.2)	4 (0.5)	66 (7.7)	6 (0.7)	14 (1.6)	1 (0.1)	2 (0.2)	1 (0.1)	13 (1.5)	28 (3.3)	
	308	15 (4.9)	6 (1.9)	77 (25)	65 (21.1)	11 (3.6)	22 (7.1)	0 (0.0)	15 (4.9)	4 (1.3)	15 (4.9)	10 (3.2)	14 (4.5)	0 (0.0)	13 (4.2)	7 (2.3)	9 (2.9)	16 (5.2)	9 (2.9)	
	173	14 (8.1)	2 (1.2)	14 (8.1)	66 (38.2)	6 (3.5)	0 (0.0)	0 (0.0)	8 (4.6)	7 (4)	5 (2.9)	25 (14.5)	9 (5.2)	0 (0.0)	2 (1.2)	4 (2.3)	1 (0.6)	3 (1.7)	7 (4)	
MTL type	1,184	387 (32.7)	149 (12.6)	93 (7.9)	164 (13.9)	15 (1.3)	31 (2.6)	13 (1.1)	37 (3.1)	43 (3.6)	24 (2.0)	94 (7.9)	23 (1.9)	14 (1.2)	12 (1.0)	12 (1.0)	7 (0.6)	29 (2.4)	37 (3.1)	
	32	13 (40.6)	0 (0.0)	1 (3.1)	9 (28.1)	3 (9.4)	1 (3.1)	1 (3.1)	0 (0.0)	0 (0.0)	0 (0.0)	2 (6.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (3.1)	
	78	31 (39.7)	9 (11.5)	13 (16.7)	4 (5.1)	1 (1.3)	2 (2.6)	0 (0.0)	6 (7.7)	2 (2.6)	0 (0.0)	5 (6.4)	3 (3.8)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.6)	
Geographical origin	226	67 (29.6)	58 (25.7)	18 (8.0)	32 (14.2)	2 (0.9)	12 (5.3)	3 (1.3)	5 (2.2)	7 (3.1)	1 (0.4)	11 (4.9)	2 (0.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	6 (2.7)	2 (0.9)	
	271	91 (33.6)	61 (22.5)	17 (6.3)	36 (13.3)	4 (1.5)	13 (4.8)	4 (1.5)	11 (4.1)	3 (1.1)	4 (1.5)	11 (4.1)	1 (0.4)	2 (0.7)	1 (0.4)	3 (1.1)	0 (0.0)	3 (1.1)	6 (2.2)	
	123	49 (39.8)	3 (2.4)	10 (8.1)	15 (12.2)	4 (3.3)	0 (0.0)	1 (0.8)	0 (0.0)	7 (5.7)	7 (5.7)	11 (8.9)	2 (1.6)	0 (0.0)	2 (1.6)	2 (1.6)	3 (2.4)	0 (0.0)	7 (5.7)	
	186	46 (24.9)	9 (4.9)	19 (10.3)	23 (12.4)	4 (2.2)	2 (1.1)	2 (1.1)	0 (0.0)	9 (4.9)	6 (3.2)	45 (24.3)	12 (6.5)	2 (1.1)	1 (0.5)	0 (0.0)	1 (0.5)	1 (0.5)	4 (2.2)	
	15	2 (13.3)	0 (0.0)	0 (0.0)	9 (60)	0 (0.0)	0 (0.0)	1 (6.7)	0 (0.0)	1 (6.7)	1 (6.7)	1 (6.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
	72	12 (18.5)	14 (21.5)	3 (4.6)	26 (40)	1 (1.5)	0 (0.0)	0 (0.0)	1 (1.5)	1 (1.5)	0 (0.0)	0 (0.0)	2 (3.1)	8 (11.1)	0 (0.0)	1 (1.5)	0 (0.0)	1 (1.5)	2 (3.1)	
	171	53 (31)	3 (1.8)	8 (4.7)	15 (8.8)	4 (2.3)	5 (2.9)	1 (0.6)	11 (6.4)	7 (4.1)	2 (1.2)	7 (4.1)	5 (2.9)	1 (0.6)	11 (6.4)	4 (2.3)	6 (3.5)	14 (8.2)	14 (8.2)	
	201	99 (49.3)	20 (10)	31 (15.4)	18 (9.0)	1 (0.5)	4 (2.0)	1 (0.5)	4 (2.0)	5 (2.5)	0 (0.0)	6 (3.0)	1 (0.5)	0 (0.0)	1 (0.5)	2 (1.0)	0 (0.0)	2 (1.0)	6 (3)	
	81	32 (39.5)	0 (0.0)	3 (3.7)	3 (3.7)	0 (0.0)	0 (0.0)	1 (1.2)	13 (16)	7 (8.6)	1 (1.2)	10 (12.3)	3 (3.7)	1 (1.2)	0 (0.0)	1 (1.2)	1 (1.2)	5 (6.2)	0 (0.0)	
	25	7 (28.0)	2 (8.0)	6 (24.0)	5 (20.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (8.0)	0 (0.0)	2 (8.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (4)	
	Anatomical origin	579	161 (27.8)	76 (13.1)	49 (8.5)	103 (17.8)	5 (0.9)	11 (1.9)	6 (1.0)	28 (4.8)	23 (4.0)	4 (0.7)	42 (7.3)	12 (2.1)	1 (0.2)	8 (1.4)	7 (1.2)	2 (0.3)	22 (3.8)	19 (3.3)
		329	118 (35.9)	54 (16.4)	22 (6.7)	32 (9.7)	8 (2.4)	18 (5.5)	4 (1.2)	5 (1.5)	12 (3.6)	2 (0.6)	23 (7.0)	10 (3.0)	0 (0.0)	1 (0.3)	1 (0.3)	3 (0.9)	6 (1.8)	10 (3)
		129	52 (40)	12 (9.2)	10 (7.7)	14 (10.8)	3 (2.3)	1 (0.8)	2 (1.5)	4 (3.1)	2 (1.5)	4 (3.1)	6 (4.6)	0 (0.0)	11 (8.5)	1 (0.8)	3 (2.3)	1 (0.8)	0 (0.0)	3 (2.3)
		214	87 (40.8)	13 (6.1)	17 (8.0)	19 (8.9)	3 (1.4)	2 (0.9)	2 (0.9)	6 (2.8)	8 (3.8)	5 (2.3)	27 (12.7)	6 (2.8)	2 (0.9)	4 (1.9)	2 (0.9)	2 (0.9)	3 (1.4)	6 (2.8)
Date of isolation	118	53 (44.9)	11 (9.3)	8 (6.8)	18 (15.3)	5 (4.2)	1 (0.8)	4 (3.4)	6 (5.1)	2 (1.7)	1 (0.8)	5 (4.2)	0 (0.0)	0 (0.0)	1 (0.8)	1 (0.8)	0 (0.0)	0 (0.0)	2 (1.7)	
	236	103 (43.6)	24 (10.2)	23 (9.7)	15 (6.4)	1 (0.4)	12 (5.1)	3 (1.3)	3 (1.3)	6 (2.5)	3 (1.3)	11 (4.7)	9 (3.8)	1 (0.4)	0 (0.0)	1 (0.4)	2 (0.8)	7 (3.0)	12 (5.1)	
	833	249 (29.9)	111 (13.3)	67 (8.0)	127 (15.2)	12 (1.4)	20 (2.4)	7 (0.8)	29 (3.5)	35 (4.2)	9 (1.1)	74 (8.9)	15 (1.8)	3 (0.4)	13 (1.6)	10 (1.2)	7 (0.8)	23 (2.8)	22 (2.6)	
Reduced susceptibility to antifungal agent ^b	48	18 (37.5)	2 (4.2)	6 (12.5)	4 (8.3)	4 (8.3)	5 (10.4)	1 (2.1)	1 (2.1)	0 (0.0)	0 (0.0)	2 (4.2)	2 (4.2)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (6.3)	
	36	12 (33.3)	2 (5.6)	1 (2.8)	3 (8.3)	3 (8.3)	3 (8.3)	1 (2.8)	1 (2.8)	1 (2.8)	0 (0.0)	3 (8.3)	1 (2.8)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (2.8)	4 (11.1)	
	17	6 (35.3)	1 (5.9)	1 (5.9)	1 (5.9)	1 (5.9)	1 (5.9)	1 (5.9)	1 (5.9)	0 (0.0)	0 (0.0)	2 (11.8)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (11.8)	
	33	24 (72.7)	1 (3.0)	0 (0.0)	0 (0.0)	1 (3.0)	2 (6.1)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (6.1)	1 (3.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (3.0)	1 (3)	

^a Includes other sterile sites, such as cerebrospinal fluid and intraoperative samples.^b Isolates were included in this category if fluconazole MICs were >8 µg/ml, itraconazole MICs were >0.25 µg/ml, voriconazole MICs were >1 µg/ml, and flucytosine MICs were >4 µg/ml. These cutoffs include isolates in the intermediate and susceptible, dose-dependent categories, as defined by the CLSI (43).

clade (Table 3), with its 104 assigned isolates constituting 7.5% of all isolates sequenced.

Characteristics of clades. To determine the statistical significance of the differentially distributed properties of isolates in the 17 clades shown in Table 3, a univariate ANOVA was run, with the clade number as the dependent variable and with ABC type, *MAT* type, geographical origin, anatomical origin, decade of isolation, and reduced susceptibility to the four antifungal agents listed in Table 3 as fixed factors. The data were analyzed with the singletons excluded and with the anatomical source omitted for isolates of animal origin. The ABC type ($P < 10^{-6}$) and geographical origin ($P = 0.005$) were highly significant factors relating to clade numbers, while the *MAT* type ($P = 0.88$), decade of isolation ($P = 0.61$), anatomical origin ($P = 0.72$), and reduced susceptibility to fluconazole ($P = 0.28$), itraconazole ($P = 0.83$), voriconazole ($P = 0.27$), and flucytosine ($P = 0.45$) did not show significant associations with clade designations.

The clades obviously differed considerably in the proportions of ABC types represented (Table 3). Clades 1 and 2 comprised >93% type A isolates: among 56 isolates of the most common clade 1 DST, DST 69, for which ABC type data were available, only a single example of type B was found. Clades 7 and 13 consisted entirely of type A isolates, and clade 9 was highly enriched with type B isolates. Clades 3, 6, 10, 14, and 16 comprised >60% isolates of type B, while type C strains showed their highest prevalence in clades 4, 5, 12, and 15.

Among 1,294 isolates typed for *MAT* status, 110 (8.5%) were homozygous, with the remaining 1,184 isolates being heterozygous. *MAT* type α/α (78 isolates) was more than twice as common as type a/a (32 isolates). Reduced susceptibilities to fluconazole, itraconazole, voriconazole, and flucytosine were all significantly associated with *MAT* homozygosity (χ^2 test; $P < 0.0001$ for each agent). Among 592 *MAT* heterozygous (a/α) isolates, 27 (4.6%), 21 (3.5%), 9 (1.5%), and 23 (3.9%) showed reduced susceptibility to fluconazole, itraconazole, voriconazole, and flucytosine, respectively. The corresponding data for 24 susceptibility-tested a/a isolates were 10 (41.7%), 5 (20.8%), 4 (16.7%), and 6 (25.0%), and those for 42 α/α isolates were 11 (26.2%), 9 (21.3%), 4 (9.5%), and 4 (9.5%). Thus, the highest prevalence of reduced susceptibility to azoles and flucytosine was found among a/a isolates, and the lowest prevalence was found among a/α isolates. None of the a/α isolates showed reduced susceptibility to all four of the agents for which resistance breakpoints were available, while four (16.8%) of the a/a isolates and two (4.8%) of the α/α isolates showed reduced susceptibility to all four agents.

Isolates in the most populous clades, clades 1 to 4, could be found from most geographical sources; nevertheless, the differential enrichment of these and other clades with isolates from different areas was statistically significant. North American isolates were predominantly assigned to clades 1 and 3 (Table 3). Among the isolates in clades 2 and 6, 70% came from the United Kingdom (70% of all UK isolates were assigned to clades 1, 2, and 4). In contrast, 59% of isolates in clade 10 originated from France and other European countries outside the United Kingdom, although these made up only 11% of all the isolates from these regions. Isolates from South-east Asian countries and Japan dominated clades 14 to 17 and the singletons, while South American isolates made up 29% of

those in clade 8. Among isolates from Africa, 40% were assigned to clade 4. Although the clade distributions of isolates from England and Scotland were generally similar, isolates from Scotland were less than half as common as those from England in clades 8 and 10 and more than twice as common in clade 9. Particular mention should be made of the isolates that constituted clade 13, which was the cluster least similar to the remainder of the isolates. Most of the clade 13 isolates were examples of a phenotypically distinct variant of *C. albicans* that was originally proposed as the separate species *Candida africana* (68). Of the 14 isolates in clade 13, plus a 15th isolate, P2246, which was more closely related to the clade 13 group than to other types but fell outside the 0.04 P distance clade cutoff, 12 were vaginal isolates and 2 came from the penis; only one clade 13 isolate was from a blood culture. Nine of the 15 isolates were of African origin (Madagascar and Angola), but 1 was from Japan, 2 were from the United Kingdom, 2 were from Germany, and 1 (the blood isolate) was from Chile. Only 2 of the 15 isolates, the one from Japan and P2246, were not of DST 182. These data confirm the group as a highly distinct *C. albicans* clade with global distribution.

Although the differential distribution among *C. albicans* clades of isolates from blood, the oropharynx, the vagina, and other sources did not reach statistical significance in the univariate ANOVA, as it had in our previous analysis of 416 isolates (63), the prevalence of oropharyngeal isolates among all isolates in clades 5 and 6 (42% and 56%, respectively) was notably high, and as already described, clade 13 comprised principally vaginal isolates.

Commensal and disease-producing isolates. Although the general linear ANOVA model showed no significant associations between anatomical sources of the 1,391 isolates and their clades in the present study, our previous analysis of MLST data for a smaller isolate set did suggest such an association (63). Schmid and colleagues previously hypothesized that one group of isolates comprising the most commonly encountered strain type by Ca3 fingerprinting represents a "general purpose" strain type with a higher propensity to cause infections than other types (55). Since 19/21 isolates we were given from this strain type corresponded to our MLST clade 1, we were interested in examining whether the proportion of isolates from clade 1 associated with disease rather than commensalism differed from that proportion in other clades.

A subset of 559 isolates was chosen from the full database of 1,391 isolates on the basis of two criteria. The first was that the isolates originated from patients in Western Europe. The second was that the raw information provided with the isolates allowed them to be assigned to one of the following three classes: commensal isolates, blood isolates, and isolates obtained from superficial *Candida* infections. Results from a single region were analyzed to minimize possible geographical bias in the data, and only the Western European isolates provided adequate numbers for the analysis. Results were analyzed for 327 European blood isolates, 177 commensal isolates (132 from the oropharynx, 35 from feces, 7 from healthy vaginas, and 3 from other superficial sites), and 55 isolates from superficial infections (37 oral isolates and 18 vaginal isolates) where the commensal or pathogenic status was known unequivocally. The clade distributions of these isolates are shown in Table 4. Isolates from clades 1, 2, 3, 4, and 11 were large

TABLE 4. Proportions of European *C. albicans* isolates from blood, commensal carriage, and superficial infections in the five most populous clades

Clade	No. (%) of blood isolates	No. (%) of commensal isolates	No. (%) of superficial infections
1	90 (27.5)	57 (32.2)	26 (47.3)
2	62 (19.0)	33 (18.6)	4 (7.3)
3	27 (8.3)	15 (8.5)	5 (9.1)
4	57 (17.4)	18 (10.2)	2 (3.6)
11	24 (7.3)	30 (16.9)	4 (7.3)
Other	67 (20.5)	24 (13.6)	14 (25.5)

enough sets to be treated as separate individual groups, with the remainder of isolates (including singletons) treated as a sixth single group. By χ^2 analysis, the clade distributions shown in Table 4 differed significantly ($P < 0.001$). Comparison of only the numbers for commensal and blood isolates also showed a highly significant difference in their clade distributions (χ^2 test; $P = 0.003$); the equivalent analysis of clade distributions of commensal and superficially pathogenic isolates gave a P value of 0.013. A χ^2 test done exclusively with the data for the largest clades (1, 2, 3, 4, and 11) showed a P value of 0.003. As shown in Table 4, major contributors to differences in distributions were related to clade 1, which contained a lower proportion of the blood isolates and a higher proportion of the isolates from superficial infections than its proportion of commensal isolates, and to clade 4, with the opposite relative distribution. However, the proportion of isolates in clade 1 that came from blood cultures was lower than the proportion in all other clades except clade 11 (Table 4). Clade 4 was particularly enriched in blood isolates.

Antifungal susceptibility and resistance. Among the 658 isolates that were tested for antifungal susceptibilities, 48 (7.3%) showed reduced susceptibility to fluconazole, 35 (5.3%) showed reduced susceptibility to itraconazole, 17 (2.6%) showed reduced susceptibility to voriconazole, and 33 (5.0%) showed reduced susceptibility to flucytosine. Among these, 26, 11, 10, and 30 isolates, respectively, met the CLSI definition of resistance to the antifungal agent concerned, with the others belonging to the “susceptible, dose-dependent” or “intermediate” category (45). There were 19 isolates with low susceptibility to fluconazole but not itraconazole or voriconazole, 13 isolates with reduced susceptibility to fluconazole plus one other azole, and 16 isolates with low susceptibility to all three. An additional seven isolates showed reduced susceptibility to itraconazole but not to fluconazole or voriconazole. Of the 55 isolates with low susceptibility to one or more azole antifungal agents, 44 were oral isolates from human immunodeficiency virus-infected patients, 6 were blood isolates, 3 were vaginal isolates, 1 was from another source, and 1 isolate was from an unstated anatomical source.

Relative to the distribution of all isolates among clades, the distribution of isolates with reduced azole susceptibility showed a higher than expected prevalence in clades 5 and 6 and in the singletons, but the difference was not statistically significant (Table 3). Among the 33 isolates with low flucytosine susceptibility, 24 (72.7%) belonged to clade 1; this most common clade accounted for 33% of isolates overall, so the

TABLE 5. Congruence analysis of genotypes of 33 *C. albicans* isolates chosen to represent the diversity of the whole population^a

Gene	–ln L for best tree	–ln L vs other genes' trees	–ln L vs 200 random trees
<i>AAT1a</i>	174.2	256.9–296.2	263.9–316.3
<i>ACC1</i>	106.6	160.7–184.7	160.6–193.9
<i>ADP1</i>	197.5	350.5–401.7	355.8–455.0
<i>MP1B</i>	209.6	420.9–522.8	391.5–515.9
<i>SYA1</i>	224.8	358.8–408.6	375.7–429.4
<i>VPS13</i>	324.6	476.7–506.0	469.6–532.0
<i>ZWF1B</i>	233.8	375.1–401.1	377.9–447.3

^a Data were generated in the same manner as that for columns 1, 6, 7, and 8 of Table 2 in reference 20.

prevalence of flucytosine-resistant isolates in this clade was twice as high as expected. This result was highly significant (χ^2 test applied to data for large clades 1 to 4 and 11 as separate groups and to other isolates as a single group; $P < 0.0001$).

Phylogenetic congruence tests. For the seven genes used for MLST analysis, a test of congruence of their phylogenetic histories was carried out. Based on a set of 33 isolates chosen as described in Materials and Methods, PAUP* was used to establish the optimum ML tree for each gene. The scores for these trees (–ln likelihood [L]) are shown in the second column of Table 5. The data for each gene in turn were then fitted to the optimum trees for each of the other six genes, and the range of scores for these is shown in the third column of Table 5. Finally, to assess the significance of these scores, the data for each gene were fitted to 200 trees of random topology (Table 5, final column). The results show that in each case, the fit of a gene's data to the trees for the other six genes was no better than the fit to random trees. This implies that the phylogenetic histories of the seven genes are not congruent.

Recombination between haplotypes. The numbers of SNPs and haplotypes computationally defined by FastPhase analysis were as follows: for *AAT1*, 22 SNPs and 50 haplotypes; for *ACC1*, 18 SNPs and 33 haplotypes; for *ADP1*, 27 SNPs and 50 haplotypes; for *MP1b*, 29 SNPs and 45 haplotypes; for *SYA1*, 23 SNPs and 67 haplotypes; for *VPS13*, 27 SNPs and 90 haplotypes; and for *ZWF1b*, 24 SNPs and 89 haplotypes. These results are not compatible with a simple linear accumulation of mutations.

To test whether recombination between different haplotypes could be a significant source of new variation, computer scripts were written to examine systematically the haplotypes for each gene from the whole data set for 1,391 isolates. The analysis identified sets of four haplotypes in which two of the haplotypes could have arisen by a single recombination of the other two. For all seven gene fragments, numerous sets of this type were found (*AAT1a*, 104; *ACC1*, 43; *ADP1*, 54; *MP1b*, 23; *SYA1*, 167; *VPS13*, 361; and *ZWF1b*, 364). To test the significance of these findings, sets of randomly generated haplotypes of the same length and diversity as the real ones were tested. Recombinant sets occurred by chance less than once per simulated population, indicating that the numbers observed in the real data are highly significant ($P < 10^{-6}$).

For each of the seven genes in turn, we analyzed the genotypes for linkage disequilibrium (LD) between each pairwise combination of SNPs within the gene (Arlequin software). All

1,391 isolates were included in this analysis. The numbers of pairwise SNP combinations and the percentages of these showing significant LD ($P < 0.05$) were as follows: for *AAT1*, 231 pairs and 31%; for *ACCI*, 153 pairs and 35%; for *ADP1*, 351 pairs and 47%; for *MP1b*, 435 pairs and 41%; for *SYA1*, 276 pairs and 50%; for *VPS13*, 351 pairs and 50%; and for *ZWF1b*, 276 pairs and 39%. The lack of complete LD within genes is consistent with the existence of recombination between haplotypes.

The analysis was repeated with only the 467 isolates in clade 1. For each gene, the number of haplotypes observed within the clade and the number of possible recombinant sets were as follows: for *AAT1*, 15 haplotypes and 10 possible recombinant combinations; for *ACCI*, 7 and 2; for *ADP1*, 10 and 0; for *MP1b*, 17 and 6; for *SYA1*, 24 and 20; for *VPS13*, 30 and 122; and for *ZWF1b*, 28 and 76. Within clade 1, the haplotype diversity was about 20 to 30% of that of the entire isolate collection, but for all genes except *ADP1*, there was still evidence for recombination between haplotypes. This suggests that the clade has not arisen by simple clonal expansion and mutation.

Evolution of eBURST clusters. The largest eBURST cluster (Fig. 1) was examined in terms of the types of changes that have occurred between linked strains, i.e., those that differed at only one of the seven fragments used for MLST, regardless of the number of SNP differences within that fragment. Approximately half of the mutational events observed could be explained by a loss of heterozygosity in the putative derived isolate relative to the putative founder; in these cases, the founder had two different haplotypes for the gene in question, and the derived isolate had two identical haplotypes. The remaining events could not be explained by this mechanism but could be a result of mutation (at one or more SNPs), mitotic recombination, or genetic exchange with another isolate.

Population genetics analysis. The putative founders of the 33 largest eBURST clusters were selected as a group of isolates that are dissimilar to each other and that represent much of the diversity of the whole set. Examination of their MLST genotypes showed that there were 52 SNPs (out of the total of 171) where the minor allele was sufficiently common that all three possible SNP genotypes were present among the 33 isolates. Of these 52 SNPs, the genotype distributions for 39 did not deviate significantly from Hardy-Weinberg equilibrium ($P > 0.05$). These 39 SNPs included at least 2 SNPs in each of the seven genes.

For each of the seven genes, one or two SNPs that showed the closest correspondence to Hardy-Weinberg equilibrium were used for an analysis of inter-SNP LD. A total of 11 SNPs were analyzed. Among the 55 inter-SNP pairs, only 7 showed significant LD ($P < 0.05$).

DISCUSSION

We analyzed the multilocus DSTs and associated properties of 1,391 *C. albicans* isolates from all regions of the world, a data set more than three times as large as any analyzed previously (63). The main findings are (i) that most of the isolates could be assigned to one of 17 homogenous clade clusters, delineated by a P distance cutoff value that retained previous clade groupings determined by MLST and by Ca3 fingerprint-

ing; (ii) that the clades differed significantly in the proportions of isolates typed as A, B, or C on the basis of an *ITS1* intron (42, 43); (iii) that the clades differed significantly in the proportions of isolates that they contained from different geographical regions of the world; (iv) that proportions of (European) isolates from blood, commensal carriage, and superficial infections differed between the most populous clades; (v) that isolates with reduced susceptibilities to azole antifungals came mainly from oral infections of AIDS patients, were highly associated with *MAT* homozygous strains, particularly *a/a* types, and were most prevalent in clades 5 and 6; (vi) that isolates with reduced flucytosine susceptibility were most commonly found (>70%) among clade 1 isolates; (vii) that the types defined by overall DSTs are not congruent with individual gene trees, which implies independent evolution of the gene fragments sequenced; and (viii) that recombination events are reflected in the haplotype analysis of the gene fragments sequenced for MLST, which suggests that the putative evolutionary histories of the isolates appear to be mixed to an extent that resembles that for a sexually reproducing species.

The epidemiological implications of surveys such as the present study are limited by the fact that the isolates typed do not represent systematically obtained material from defined sources. However, the number and diversity of the 1,391 isolates represented in the present test panel give us reason for confidence in the main conclusions of the study. We have divided the isolates into 17 clades in a way that leaves a small minority of isolates characterized as singletons. The previous analysis of MLST data for 416 *C. albicans* isolates designated 12 clades (63). As the database of sequence data grows, there will be an inevitable consolidation of information to define groups of *C. albicans* clades. The point at which we have set the boundary for clade differentiation is slightly broader for the present large data set than that for the previous analysis. While all but 5 isolates assigned to 1 of 12 clades in the previous analysis retained their clade assignments, the new analysis resulted in 45 isolates previously designated singletons becoming members of definable clades. Notable among these isolates is WO-1, the original white-opaque switching isolate (2), which is the second *C. albicans* strain to have its genome sequenced (http://www.broad.mit.edu/annotation/genome/candida_albicans/Home.html). WO-1 now emerges as a member of clade 6, unlike SC5314, the first strain used for *C. albicans* genome sequencing, which belongs to clade 1 (30).

Clade 1 is the most populous strain group in this and the previous study and correlates with the most populous cluster observed in various studies based on DNA fingerprinting and other genome-based technologies (39, 49, 55, 61). This demonstrates that the most common, globally distributed *C. albicans* strain types are those of MLST clade 1, and DST 69 is both the most frequently encountered member of clade 1 and its putative founding type. The notion has been advanced that *C. albicans* clade 1 strains may have a higher propensity than other types to undergo the change from commensal to pathogen, notwithstanding the requirement for reduction in host defenses needed to permit the change of status (55). It seems reasonable to hypothesize that clade 1 strains have a high capability for spread among the human population, since examples of DST 69 alone were isolated in North and South America, Japan, Hong Kong, Malaysia, the Middle East, and

New Zealand as well as from the United Kingdom and Europe, which were the geographical sources of most of our isolates. It is less evident that clade 1 isolates have an enhanced ability to cause infection relative to other types. In our analysis limited to European isolates to minimize geographical influence, the proportion of bloodstream isolates assigned to clade 1 was smaller than the proportion of commensal isolates; however, a larger proportion of isolates from superficial infections was found in clade 1 (Table 4). In contrast, clade 4 could be regarded as particularly rich in blood isolates. Statistical analysis of the proportions of commensal and invasive isolates in different clades, as shown in Table 4, gave significant outcomes, even when subsets of the data were analyzed separately. We interpret these findings as indicative of a possible clade-related differential potential for strains to be encountered as commensals, superficial pathogens, and agents of disseminated infection. The notion that clade 1 isolates are very widely distributed and are particularly commonly found as commensals and superficial pathogens, but less often as pathogens of deep tissues, is compatible with the notion that these strains are better adapted to colonize and invade epithelial surfaces, whereas bloodstream dissemination is more likely to be a consequence of severe host immunosuppression than of virulence properties intrinsic to the fungus.

Geographical differences among *C. albicans* strain types are to be expected, since evolutionary changes from founding types should mainly have occurred independently in separated regions. However, the term "geographical enrichment" of different strain types adequately summarizes the findings with *C. albicans*. Because the species exists primarily as a commensal of humans and other warm-blooded animals, the extent to which populations remain geographically independent will always be limited by movements of the hosts. Thus, in our study, clade 2 was particularly rich in isolates from the United Kingdom, which made up 70% of the clade; more than 20% of isolates in clades 1 and 3 came from North America; isolates from the Far East were predominant among those in clades 14 to 17; and so on. A study of isolates from Taiwan by MLST clearly demonstrated differences between these and European isolates (11), but the extent of the distinction becomes diluted when these data are included in a larger data set such as the one presented here. Thus, no single clade could be described as comprised uniquely of isolates from a single geographical region. Prior work based on ABC subtyping alone found geographical variations in the prevalence of types A, B, and C (42), and this observation can be amplified by the knowledge that many MLST clades are composed mainly of isolates of the A or B type (Table 3). Care is needed in interpretation of clade geographical enrichments, since neither this nor previous studies demonstrating the effect have used panels of isolates matched for all properties except geographical origin and may therefore be geographically biased. Ca3 fingerprinting identified a clade originally described as an African clade (5), but the isolate panel tested in that study included mainly isolates from Africa. African clade isolates cluster in clade 4 by MLST. Among the 182 isolates of clade 4 in the present study, 14.3% originated from Africans and 37.4% originated from the United Kingdom, which suggests that clade 4 should not be described as Africa specific. However, the 14.7% prevalence of African isolates in clade 4 was higher than the prevalence of

African isolates in the other four most populous clades (1 to 3 and 11), confirming at least an African enrichment of this clade.

A more distinct specificity towards the African continent is also seen in clade 13, where 8 (57%) of 14 isolates were of African origin. These isolates represent the clade once proposed as the new species *C. africana* (68). Clade 13 isolates formed a cluster well separated from the other types (Fig. 1). It is notable that isolate P2246 (ABC type B, DST 797) differed substantially from the ABC type A, DST 182 or 782 that characterized the 14 isolates in clade 13, but still clustered closely with clade 13 by UPGMA. This isolate, previously numbered AM335, also differed notably from other atypical African isolates in a prior study based on a combination of molecular phylogenetic methodologies (22).

Our findings concerning antifungal resistance among the *C. albicans* isolates are not novel: the strong association of flucytosine resistance with clade 1 has been described previously (17, 63), as has the association of flucytosine and azole resistance with *MAT* homozygosity (51, 53, 63). The association has been explained in terms of homozygosity of a particular allele of *TAC1*, which upregulates *C. albicans* efflux pumps and, like the *MAT* locus, is located on chromosome 5 (13). The present study does, however, suggest that the propensity towards antifungal resistance is higher among *a/a* isolates than among α/α isolates, even though the latter type is more common in the whole population. Our data also serve as a reminder that the majority of *C. albicans* isolates showing resistance to azole antifungal agents were obtained from oral samples from HIV/AIDS patients in the 1990s. The overall prevalence of antifungal-resistant isolates can only be estimated from a random sample of clinical isolates, not from a partly selected isolate panel such as the one used in this study.

Our analysis of haplotypes for putative recombination events and the lack of phylogenetic congruence between the seven gene fragments sequenced point to apparently separate evolutionary histories for these seven genes, with recombination events intervening sufficiently often to remove what would otherwise be expected to be a generally clonal evolutionary pattern in an asexually reproducing species. Mitotic recombination between homologous chromosomes, followed by parasexual chromosome exchange, could explain the high level of generation of new haplotypes. This scenario is difficult, if not impossible, to distinguish from true sexual reproduction, since either process could result in similar outcomes. Since chromosomal loss and reduplication events are already well evidenced in *C. albicans* (31, 57, 71, 72), it will be interesting to test examples of strain pairs differing only by a loss of heterozygosity at one MLST locus for complete or partial aneuploidy by sequencing markers across whole chromosomes or by comparative genome hybridization (57). The possibility of cryptic mating events may also contribute to the production of a population in which the most diverse strains are to a large extent in Hardy-Weinberg equilibrium, with a lack of LD between loci. In this respect, *C. albicans* resembles a sexually reproducing species (16, 26, 33, 36, 59, 60). Superimposed on these infrequent genetic recombination events are many small incremental changes, such as losses of heterozygosity and accumulations of single mutations, that can be seen by examining the relationships of strains within eBURST clusters.

These genetic events are the most probable underpinnings of the tendency of *C. albicans* isolates to be highly similar, but not always indistinguishable, among members of families (8) and among sequential samples from individuals (47). The maintenance of high levels of genetic diversity in the absence of frequent sexual matings may characterize success for survival of a fungus so widespread as a commensal that becomes a pathogen only when host antimicrobial defenses are impaired.

ACKNOWLEDGMENTS

This work is based on data generated from several projects that were supported by the Wellcome Trust (grants 069615 and 074898), the Ministère de la Recherche et de la Technologie (Programme de Recherche Fondamentale en Microbiologie, Maladies Infectieuses et Parasitaires-Réseau Infections Fongiques), Pasteur Genopole-Ile-de-France, Merck, Sharp & Dohme, Pfizer UK, the Center for Disease Control, Department of Health, Taiwan (grant DOH94-DC-2011), the National Science Council, Taiwan (grant 94-0324-19-F-00-00-00-35), and the European Commission (Euresfun; grant LSHM-CT-2005-518199).

We are grateful to the many individuals who have supplied us with the *C. albicans* isolates, past and present, that formed the experimental material for this study, including J. B. Anderson, G. Chaves, M. Cuenca-Estrella, D. H. Ellis, J.-M. Gomez, G. Haase, M. Hanson, R. Hollis, E. M. Johnson, B. Jones, G. Just, C. C. Kibbler, N. Noland, M. A. Pfaller, J.-L. Rodriguez-Tudela, J. Schmid, D. R. Soll, D. A. Stevens, F. Symoens, S. Takakura, and K. Y. Yuen. We gratefully acknowledge the skilled technical assistance of Christiane Bouchier, Diana Sharafi, and Julie Whyte. Keith Jolley (University of Oxford) provided valuable guidance in carrying out congruence analysis.

REFERENCES

- Anderson, J. B., C. Wickens, M. Khan, L. E. Cowen, N. Federspiel, T. Jones, and L. M. Kohn. 2001. Infrequent genetic exchange and recombination in the mitochondrial genome of *Candida albicans*. *J. Bacteriol.* **183**:865–872.
- Anderson, J. M., and D. R. Soll. 1987. The unique phenotype of opaque cells in the “white-opaque transition” in *Candida albicans*. *J. Bacteriol.* **169**:5579–5588.
- Bennett, R. J., and A. D. Johnson. 2003. Completion of a parasexual cycle in *Candida albicans* by induced chromosome loss in tetraploid strains. *EMBO J.* **22**:2505–2515.
- Bennett, R. J., and A. D. Johnson. 2005. Mating in *Candida albicans* and the search for a sexual cycle. *Annu. Rev. Microbiol.* **59**:233–255.
- Blignaut, E., C. Pujol, S. Lockhart, S. Joly, and D. R. Soll. 2002. Ca3 fingerprinting of *Candida albicans* isolates from human immunodeficiency virus-positive and healthy individuals reveals a new clade in South Africa. *J. Clin. Microbiol.* **40**:826–836.
- Bougnoux, M.-E., D. M. Aanensen, S. Morand, M. Théraud, B. G. Spratt, and C. d’Enfert. 2004. Multilocus sequence typing of *Candida albicans*: strategies, data exchange and applications. *Infect. Genet. Evol.* **4**:243–252.
- Bougnoux, M.-E., S. Morand, and C. d’Enfert. 2002. Usefulness of multilocus sequence typing for characterization of clinical isolates of *Candida albicans*. *J. Clin. Microbiol.* **40**:1290–1297.
- Bougnoux, M. E., D. Diogo, N. Francois, B. Sendid, S. Veirmeire, J. F. Colombel, C. Bouchier, H. Van Kruiningen, C. d’Enfert, and D. Poulain. 2006. Multilocus sequence typing reveals intrafamilial transmission and microevolutions of *Candida albicans* isolates from the human digestive tract. *J. Clin. Microbiol.* **44**:1810–1820.
- Bougnoux, M. E., A. Tavanti, C. Bouchier, N. A. R. Gow, A. Magnier, A. D. Davidson, M. C. J. Maiden, C. d’Enfert, and F. C. Odds. 2003. Collaborative consensus for optimized multilocus sequence typing of *Candida albicans*. *J. Clin. Microbiol.* **41**:5265–5266.
- Calderone, R. A. 2002. *Candida* and candidiasis. ASM Press, Washington, DC.
- Chen, K. W., Y. C. Chen, H. J. Lo, F. C. Odds, T. H. Wang, C. Y. Lin, and S. Y. Li. 2006. Multilocus sequence typing for analyses of clonality of *Candida albicans* strains in Taiwan. *J. Clin. Microbiol.* **44**:2172–2178.
- Chibana, H., J. L. Beckerman, and P. T. Magee. 2000. Fine-resolution physical mapping of genomic diversity in *Candida albicans*. *Genome Res.* **10**:1865–1877.
- Coste, A., V. Turner, F. Ischer, J. Morschhauser, A. Forche, A. Selmecki, J. Berman, J. Bille, and D. Sanglard. 2006. A mutation in Tac1p, a transcription factor regulating *CDR1* and *CDR2*, is coupled with loss of heterozygosity at chromosome 5 to mediate antifungal resistance in *Candida albicans*. *Genetics* **172**:2139–2156.
- Cuenca-Estrella, M., A. Gomez-Lopez, E. Mellado, and J. L. Rodriguez-Tudela. 2005. Correlation between the procedure for antifungal susceptibility testing for *Candida* spp. of the European Committee on Antibiotic Susceptibility Testing (EUCAST) and four commercial techniques. *Clin. Microbiol. Infect.* **11**:486–492.
- Cuenca-Estrella, M., C. B. Moore, F. Barchiesi, J. Bille, E. Chrysanthou, D. W. Denning, J. P. Donnelly, F. Dromer, B. Dupont, J. H. Rex, M. D. Richardson, B. Sancak, P. E. Verweij, and J. L. Rodriguez-Tudela. 2003. Multicenter evaluation of the reproducibility of the proposed antifungal susceptibility testing method for fermentative yeasts of the Antifungal Susceptibility Testing Subcommittee of the European Committee on Antimicrobial Susceptibility Testing (AFST-EUCAST). *Clin. Microbiol. Infect.* **9**:467–474.
- Daniels, K. J., T. Srikantha, S. R. Lockhart, C. Pujol, and D. R. Soll. 2006. Opaque cells signal white cells to form biofilms in *Candida albicans*. *EMBO J.* **25**:2240–2252.
- Dodgson, A. R., K. J. Dodgson, C. Pujol, M. A. Pfaller, and D. R. Soll. 2004. Clade-specific flucytosine resistance is due to a single nucleotide change in the *FUR1* gene of *Candida albicans*. *Antimicrob. Agents Chemother.* **48**:2223–2227.
- Dodgson, A. R., C. Pujol, D. W. Denning, D. R. Soll, and A. J. Fox. 2003. Multilocus sequence typing of *Candida glabrata* reveals geographically enriched clades. *J. Clin. Microbiol.* **41**:5709–5717.
- Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform.* **1**:47–50.
- Feil, E. J., E. C. Holmes, D. E. Bessen, M.-S. Chan, N. P. J. Dayi, M. C. Enright, R. Goldstein, D. W. Hood, A. Kalai, C. E. Moore, J. Zhou, and B. G. Spratt. 2001. Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. *Proc. Natl. Acad. Sci. USA* **98**:182–187.
- Feil, E. J., B. C. Li, D. M. Aanensen, W. P. Hanage, and B. G. Spratt. 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J. Bacteriol.* **186**:1518–1530.
- Forche, A., G. Schonian, Y. Gräser, R. Vilgalys, and T. G. Mitchell. 1999. Genetic structure of typical and atypical populations of *Candida albicans* from Africa. *Fungal Genet. Biol.* **28**:107–125.
- Gräser, Y., M. Volovsek, J. Arrington, G. Schonian, W. Presber, T. G. Mitchell, and R. Vilgalys. 1996. Molecular markers reveal that population structure of the human pathogen *Candida albicans* exhibits both clonality and recombination. *Proc. Natl. Acad. Sci. USA* **93**:12473–12477.
- Holmes, E. C., R. Urwin, and M. J. Maiden. 1999. The influence of recombination on the population structure and evolution of the human pathogen *Neisseria meningitidis*. *Mol. Biol. Evol.* **16**:741–749.
- Hull, C. M., and A. D. Johnson. 1999. Identification of a mating type-like locus in the asexual pathogenic yeast *Candida albicans*. *Science* **285**:1271–1275.
- Hull, C. M., R. M. Raisner, and A. D. Johnson. 2000. Evidence for mating of the “asexual” yeast *Candida albicans* in a mammalian host. *Science* **289**:307–310.
- Hunter, P. R. 1990. Reproducibility and indices of discriminatory power of microbial typing methods. *J. Clin. Microbiol.* **28**:1903–1905.
- Jacobsen, M. D., N. A. R. Gow, M. C. J. Maiden, D. J. Shaw, and F. C. Odds. 2007. Strain typing and determination of population structure of *Candida krusei* by multilocus sequence typing. *J. Clin. Microbiol.* **45**:317–323.
- Johnson, A. 2003. The biology of mating in *Candida albicans*. *Nat. Rev. Microbiol.* **1**:106–116.
- Jones, T., N. A. Federspiel, H. Chibana, J. Dungan, S. Kalman, B. B. Magee, G. Newport, Y. R. Thorstenson, N. Agabian, P. T. Magee, R. W. Davis, and S. Scherer. 2004. The diploid genome sequence of *Candida albicans*. *Proc. Natl. Acad. Sci. USA* **101**:7329–7334.
- Kabir, M. A., A. Ahmad, J. R. Greenberg, Y. K. Wang, and E. Rustchenko. 2005. Loss and gain of chromosome 5 controls growth of *Candida albicans* on sorbose due to dispersed redundant negative regulators. *Proc. Natl. Acad. Sci. USA* **102**:12147–12152.
- Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* **5**:150–163.
- Lachke, S. A., S. R. Lockhart, K. J. Daniels, and D. R. Soll. 2003. Skin facilitates *Candida albicans* mating. *Infect. Immun.* **71**:4970–4976.
- Lan, C. Y., G. Newport, L. A. Murillo, T. Jones, S. Scherer, R. W. Davis, and N. Agabian. 2002. Metabolic specialization associated with phenotypic switching in *Candida albicans*. *Proc. Natl. Acad. Sci. USA* **99**:14907–14912.
- Legrand, M., P. Lephart, A. Forche, F. M. C. Mueller, T. Walsh, P. T. Magee, and B. B. Magee. 2004. Homozygosity at the MTL locus in clinical strains of *Candida albicans*: karyotypic rearrangements and tetraploid formation. *Mol. Microbiol.* **52**:1451–1462.
- Lockhart, S. R., K. J. Daniels, R. Zhao, D. Wessels, and D. R. Soll. 2003. Cell biology of mating in *Candida albicans*. *Eukaryot. Cell* **2**:49–61.
- Lockhart, S. R., J. J. Fritch, A. S. Meier, K. Schroppel, T. Srikantha, R. Galask, and D. R. Soll. 1995. Colonizing populations of *Candida albicans* are

- clonal in origin but undergo microevolution through C1 fragment reorganization as demonstrated by DNA fingerprinting and C1 sequencing. *J. Clin. Microbiol.* **33**:1501–1509.
38. Lockhart, S. R., C. Pujol, K. J. Daniels, M. G. Miller, A. D. Johnson, M. A. Pfaller, and D. R. Soll. 2002. In *Candida albicans*, white-opaque switchers are homozygous for mating type. *Genetics* **162**:737–745.
 39. Lott, T. J., B. P. Holloway, D. A. Logan, R. Fundyga, and J. Arnold. 1999. Towards understanding the evolution of the human commensal yeast *Candida albicans*. *Microbiology* **145**:1137–1143.
 40. Maiden, M. C. J. 2006. Multilocus sequence typing of bacteria. *Annu. Rev. Microbiol.* **60**:561–588.
 41. Mata, A. L., R. T. Rosa, E. A. R. Rosa, R. B. Goncalves, and J. F. Hoffing. 2000. Clonal variability among oral *Candida albicans* assessed by allozyme electrophoresis analysis. *Oral Microbiol. Immunol.* **15**:350–354.
 42. McCullough, M., K. V. Clemons, and D. A. Stevens. 1999. Molecular epidemiology of the global and temporal diversity of *Candida albicans*. *Clin. Infect. Dis.* **29**:1220–1225.
 43. McCullough, M. J., K. V. Clemons, and D. A. Stevens. 1999. Molecular and phenotypic characterization of genotypic *Candida albicans* subgroups and comparison with *Candida dubliniensis* and *Candida stellatoidea*. *J. Clin. Microbiol.* **37**:417–421.
 44. Miller, M. G., and A. D. Johnson. 2002. White-opaque switching in *Candida albicans* is controlled by mating-type locus homeodomain proteins and allows efficient mating. *Cell* **110**:293–302.
 45. NCCLS. 2002. Reference method for broth dilution antifungal susceptibility testing of yeasts. Approved standard—2nd ed., vol. 17, p. 9. NCCLS, Wayne, PA.
 46. Odds, F. C. 1988. *Candida and candidosis* (2nd ed.). Bailliere Tindall, London, United Kingdom.
 47. Odds, F. C., A. D. Davidson, M. D. Jacobsen, A. Tavanti, J. A. Whyte, C. C. Kibbler, D. H. Ellis, M. C. J. Maiden, D. J. Shaw, and N. A. R. Gow. 2006. *Candida albicans* strain maintenance, replacement, and microvariation demonstrated by multilocus sequence typing. *J. Clin. Microbiol.* **44**:3647–3658.
 48. Poulter, R. T. 1987. Natural auxotrophic heterozygosity in *Candida albicans*. *Crit. Rev. Microbiol.* **15**:97–101.
 49. Pujol, C., S. Joly, S. R. Lockhart, S. Noel, M. Tibayrenc, and D. R. Soll. 1997. Parity among the randomly amplified polymorphic DNA method, multilocus enzyme electrophoresis, and Southern blot hybridization with the moderately repetitive DNA probe Ca3 for fingerprinting *Candida albicans*. *J. Clin. Microbiol.* **35**:2348–2358.
 50. Pujol, C., S. Joly, B. Nolan, T. Srikantha, and D. R. Soll. 1999. Microevolutionary changes in *Candida albicans* identified by the complex Ca3 fingerprinting probe involve insertions and deletions of the full-length repetitive sequence RPS at specific genomic sites. *Microbiology* **145**:2635–2646.
 51. Pujol, C., S. A. Messer, M. Pfaller, and D. R. Soll. 2003. Drug resistance is not directly affected by mating type locus zygosity in *Candida albicans*. *Antimicrob. Agents Chemother.* **47**:1207–1212.
 52. Pujol, C., J. Reynes, F. Renaud, M. Raymond, M. Tibayrenc, F. J. Ayala, F. Janbon, M. Mallie, and J. M. Bastide. 1993. The yeast *Candida albicans* has a clonal mode of reproduction in a population of infected human immunodeficiency virus-positive patients. *Proc. Natl. Acad. Sci. USA* **90**:9456–9459.
 53. Rustad, T. R., D. A. Stevens, M. A. Pfaller, and T. C. White. 2002. Homozygosity at the *Candida albicans* MTL locus associated with azole resistance. *Microbiology* **148**:1061–1072.
 54. Scheet, P., and M. Stephens. 2006. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* **78**:629–644.
 55. Schmid, J., S. Herd, P. R. Hunter, R. D. Cannon, M. S. M. Yasin, S. Samad, M. Carr, D. Parr, W. McKinney, M. Schoushoe, B. Harris, R. Ikram, M. Harris, A. Restrepo, G. Hoyos, and K. P. Singh. 1999. Evidence for a general-purpose genotype in *Candida albicans*, highly prevalent in multiple geographical regions, patient types and types of infection. *Microbiology* **145**:2405–2413.
 56. Schröppel, K., M. Rotman, R. Galask, K. Mac, and D. R. Soll. 1994. Evolution and replacement of *Candida albicans* strains during recurrent vaginitis demonstrated by DNA fingerprinting. *J. Clin. Microbiol.* **32**:2646–2654.
 57. Selmecki, A., S. Bergmann, and J. Berman. 2005. Comparative genome hybridization reveals widespread aneuploidy in *Candida albicans* laboratory strains. *Mol. Microbiol.* **55**:1553–1565.
 58. Selmecki, A., A. Forche, and J. Berman. 2006. Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science* **313**:367–370.
 59. Soll, D. R. 2004. Mating-type locus homozygosity, phenotypic switching and mating: a unique sequence of dependencies in *Candida albicans*. *Bioessays* **26**:10–20.
 60. Soll, D. R., S. R. Lockhart, and R. Zhao. 2003. Relationship between switching and mating in *Candida albicans*. *Eukaryot. Cell* **2**:390–397.
 61. Soll, D. R., and C. Pujol. 2003. *Candida albicans* clades. *FEMS Immunol. Med. Microbiol.* **39**:1–7.
 62. Spratt, B. G., W. P. Hanage, B. Li, D. M. Aanensen, and E. J. Feil. 2004. Displaying the relatedness among isolates of bacterial species—the eBURST approach. *FEMS Microbiol. Lett.* **241**:129–134.
 63. Tavanti, A., A. D. Davidson, M. J. Fordyce, N. A. R. Gow, M. C. J. Maiden, and F. C. Odds. 2005. Population structure and properties of *Candida albicans*, as determined by multilocus sequence typing. *J. Clin. Microbiol.* **43**:5601–5613.
 64. Tavanti, A., A. D. Davidson, E. M. Johnson, M. C. J. Maiden, D. J. Shaw, N. A. R. Gow, and F. C. Odds. 2005. Multilocus sequence typing for differentiation of strains of *Candida tropicalis*. *J. Clin. Microbiol.* **43**:5593–5600.
 65. Tavanti, A., N. A. R. Gow, M. C. J. Maiden, F. C. Odds, and D. J. Shaw. 2004. Genetic evidence for recombination in *Candida albicans* based on haplotype analysis. *Fungal Genet. Biol.* **41**:553–562.
 66. Tavanti, A., N. A. R. Gow, S. Senesi, M. C. J. Maiden, and F. C. Odds. 2003. Optimization and validation of multilocus sequence typing for *Candida albicans*. *J. Clin. Microbiol.* **41**:3765–3776.
 67. Taylor, J. W., and M. C. Fisher. 2003. Fungal multilocus sequence typing—it's not just for bacteria. *Curr. Opin. Microbiol.* **6**:351–356.
 68. Tietz, H. J., M. Hopp, A. Schmalreck, W. Sterry, and V. Czaika. 2001. *Candida africana* sp. nov., a new human pathogen or a variant of *Candida albicans*? *Mycoses* **44**:437–445.
 69. Tsong, A. E., M. G. Miller, R. M. Raisner, and A. D. Johnson. 2003. Evolution of a combinatorial transcriptional circuit: a case study in yeasts. *Cell* **115**:389–399.
 70. Tzung, K. W., R. M. Williams, S. Scherer, N. Federspiel, T. Jones, N. Hansen, V. Bivolarevic, L. Huizar, C. Komp, R. Surzycki, R. Tamse, R. W. Davis, and N. Agabian. 2001. Genomic evidence for a complete sexual cycle in *Candida albicans*. *Proc. Natl. Acad. Sci. USA* **98**:3249–3253.
 71. Wellington, M., and E. Rustchenko. 2005. 5-Fluoro-orotic acid induces chromosome alterations in *Candida albicans*. *Yeast* **22**:57–70.
 72. Wu, W., C. Pujol, S. R. Lockhart, and D. R. Soll. 2005. Chromosome loss followed by duplication is the major mechanism of spontaneous mating-type locus homozygosity in *Candida albicans*. *Genetics* **169**:1311–1327.